View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by UBA - FCE (E-Journals - Facultad de Ciencias Económicas, Universidad de Buenos Aires)

RAZONAMIENTO REVOCABLE Y LÓGICAS NO MONÓTONAS: UN ANÁLISIS CONCEPTUAL[1]

Javier Legris

Se llama *revocables* a todos aquellos razonamientos no deductivos en los cuales la conclusión es obtenida a partir de información incompleta, es decir, a partir de premisas que representan condiciones insuficientes para afirmar la conclusión. En los razonamientos revocables, nueva información que se agregue a las premisas puede invalidar la inferencia hecha, pues daría lugar a inconsistencias. Así, la relación de inferencia que subyace a los razonamientos revocables no cumple con la propiedad de monotonía que satisface la relación usual de inferencia deductiva. Las lógicas no monótonas han surgido con la idea de formalizar este tipo de razonamientos. En este trabajo se sostendrá que el problema de la revocabilidad no atañe a operadores lógicos ni a la lógica en sentido estricto, sino a cómo la lógica *es aplicada* en diferentes contextos, lo que implica tomar en cuenta factores externos a la caracterización de operadores lógicos. Esta perspectiva no conduce a una modificación de la lógica, sino de la manera en que se organiza la información a partir de la cual se hacen las inferencias. Así, se debe *agregar estructura* a la información, considerando un razonamiento revocable como una estructura *más compleja* que un razonamiento deductivo.

[1] Este trabajo refleja algunos de los resultados obtenidos en el marco de los proyectos de investigación EC001/J y JE01 dirigidos por el autor y subvencionados por la Secretaría de Ciencia y Técnica de la UBA. El trabajo debe mucho a las discusiones sostenidas por el autor con Silvia Lerner y Carlos Lombardi durante varios años en el marco de estos proyectos. Por cierto, ellos no son responsables por los errores que puedan quedar.

1. INTRODUCCIÓN: EL CONCEPTO DE RAZONAMIENTO REVOCABLE

Se llama *revocables*[2] a todos aquellos razonamientos no deductivos en los cuales la conclusión es inferida a partir de información incompleta, es decir, a partir de premisas que representan condiciones insuficientes para afirmar la conclusión. Razonamientos de este tipo son habituales en un gran número de contextos tales como la investigación científica, el diseño de artefactos de todo tipo, la planificación de organismos sociales, el diagnóstico de situaciones, etc.

Los siguientes dos ejemplos (mencionados en la bibliografía sobre el tema) servirán para dar una primera idea de este tipo de razonamientos.

(1) Si el termostato está puesto en 21 C y en la habitación hace 25 C, entonces el acondicionador de aire lanzará aire frío.
El termostato está puesto en 21 C y en la habitación hace 25 C.
Luego, el acondicionador de aire lanzará aire frío.

(2) La mayoría de las aves vuelan.
Tweety es un ave.
Luego, Tweety vuela

---

[2] En este trabajo la expresión "razonamiento revocable" es la traducción del inglés *defeasible reasoning*, que a veces se traduce también como "razonamiento rebatible", e incluso como "razonamiento derrotable". La palabra "derrotable" da una idea de competencia o "lucha" entre razonamientos, lo que no siempre es el caso. "revocable" comparte con *defeasible* un matiz jurídico ("revocar leyes"), ausente en "rebatible".

El razonamiento (1) es revocable, pues no están indicadas *todas* las condiciones suficientes para la conclusión, tales como que el acondicionador de aire esté conectado, que no haya un corte de energía eléctrica, etc. En (2), hablar de “la mayoría” de las aves *no es suficiente* para extraer conclusiones acerca de un pájaro en particular.

En los razonamientos revocables, nueva información que se agregue a las premisas puede invalidar la inferencia hecha, pues daría lugar a inconsistencias. Piénsese en el primer ejemplo. Si se agrega como premisas adicionales:

- Si hay un corte de energía eléctrica, entonces el acondicionador de aire no lanzará aire frío.
- Hay un corte de energía eléctrica.
  Entonces puede inferirse deductivamente
- El acondicionador de aire no lanzará aire frío,
  lo que contradice la conclusión de (1). Respecto del segundo ejemplo, si se agregara a las premisas el enunciado
- Tweety es un pingüino,
  sabiendo además que
- Ningún pingüino vuela,
  se infiere, entonces
- Tweety no vuela,
  lo que contradice la conclusión de (2).

Es así que la relación de inferencia que subyace a los razonamientos revocables no cumple con la propiedad de *monotonía* característica de la relación usual de inferencia deductiva. Si se representa mediante el signo $\vdash$ la relación de inferencia deductiva, la propiedad de monotonía dice:

$$(\mathit{Mon}) \text{ Si } A_1, .., A_n \vdash B, \text{ entonces } C, A_1, .., A_n \vdash B,$$

donde $A_1, .., A_n, B$ y C representan enunciados cualesquiera. En otras palabras, la propiedad de monotonía dice que el agregado de premisas no modifica la validez de la inferencia.[3] En los razonamientos recién expuestos, no se cumple esta propiedad. El análisis formal de este tipo de casos ha dado origen a las *lógicas no monótonas*.

2. RAZONAMIENTOS REVOCABLES Y LÓGICA APLICADA

La *inteligencia artificial* se interesó desde sus comienzos por desarrollar teorías acerca del comportamiento de agentes que deben extraer conclusiones a partir de la información dada y, llegado el caso, tomar decisiones a partir de estas conclusiones. Por esta razón, entre sus temas de investigación se encuentra el de las inferencias revocables. Como dice Raymond Reiter:

"Si los investigadores en Inteligencia Artificial (IA) pueden estar de acuerdo en algo, es que un artefacto inteligente debe ser capaz de razonar acerca del mundo que habita. Este artefacto debe poseer varias formas de conocimiento y creencias acerca de su mundo, y debe usar esa información para inferir información ulterior acerca de aquel mundo con el fin de tomar decisiones, planificar y realizar

---

[3] Otras dos propiedades esenciales de la deducción son la reflexividad y la transitividad, expresadas respectivamente mediante los siguientes principios:

(Refl) $A \vdash A$

(Trans) Si $A_1, .., A_m \vdash B$ y $B, C_1, .., C_n \vdash D$, entonces $A_1, .., A_m, C_1, .., C_n \vdash D$.

acciones, responder a otros agentes, etc. El problema técnico para la IA es caracterizar los patrones de razonamientos que requiere semejante artefacto inteligente y realizarlos computacionalmente.” (Reiter 1987, p. 147)

Existen diversos enfoques para analizar estos patrones y entre ellos está el de las lógicas no monótonas[4]. De este modo, “el campo del razonamiento no monótono es la derivación de conclusiones plausibles (pero no infalibles) a partir de una base de conocimiento considerada abstractamente como un conjunto de fórmulas en una lógica conveniente. Cualquier conclusión de este tipo se entiende como tentativa; puede ser que se retracte de ella luego de que nueva información haya sido añadida a la base de conocimiento.” (Reiter 1987, p. 148).

La investigación *lógica* del razonamiento revocable ha sido muy fructífera y ha ampliado los horizontes e intereses de la lógica matemática en las últimas dos décadas, originando todo un campo de aplicación de la lógica a la modelización del conocimiento y la acción humanos, en el que se han obtenido importantes logros. Un enorme número de sistemas se han desarrollado sobre la base de diferentes puntos de vista, y la discusión ha avanzado notablemente a través de un gran número de reuniones científicas regulares en diversas partes del mundo y de una larga serie de revistas especializadas en este tema, tales como *Journal of Logic and Computation*, *Journal of Applied Non-Classical Logics*, *Journal of the*

[4] Otros enfoques son, por ejemplo, el probabilístico y el de la lógica borrosa.

Interest Group in Pure and Applied Logic, esto sin considerar las publicaciones periódicas dedicadas a la lógica matemática y la inteligencia artificial y a la epistemología. El futuro de esta área de investigación es promisorio.

Algunos de los enfoques más importantes son (Brewka, 1991): lógica no monótona modal, lógica autoepistémica, lógica *default*, circunscripción, razonamiento hipotético, lógicas de condicionales revocables, sistemas basados en reglas. A estos enfoques deben agregarse otras teorías emparentadas con el razonamiento revocable como la argumentación rebatible y la teoría de cambio de creencias.

La posibilidad de aplicación de estos formalismos a sistemas de información en administración es mencionada con frecuencia en tratados del área, (por ejemplo Poole, Mackworth & Goebel, 1998) y se han construido algunos ejemplos. Los teóricos de la racionalidad acotada han tomado algunos resultados de la inteligencia artificial, de la teoría de la decisión y de la teoría de juegos y han propuesto algunas implementaciones computacionales. Análogamente, la interacción con el área de razonamiento revocable es prometedora.[5]

[5] La concepción de la racionalidad acotada iniciada por Herbert Simon en sus estudios sobre administración constituye una aproximación para analizar los fenómenos económicos que se refiere a la elección racional que toma en consideración las limitaciones cognitivas del decisor, limitaciones que pueden vincularse tanto con el conocimiento como con la capacidad computacional (véase, por ejemplo, Simon 1981). Las limitaciones afectan la capacidad de hacer inferencias bajo incertidumbre. Los que investigan acerca de la racionalidad acotada coinciden en que hay supuestos del modelo neoclásico en economía que deben ser revisados, como los de conocimiento perfecto y el de procesamiento óptimo de la información.

3. REVOCABILIDAD Y CONSTANTES LÓGICAS

En los razonamientos revocables aparecen, como un elemento esencial de los mismos, enunciados "por omisión" o "de normalidad" llamados usualmente *defaults*. Estos incluyen expresiones como "la mayoría de", "normalmente", "generalmente", "típicamente", etc. Así, el problema de determinar la naturaleza de la inferencia revocable pasa por reconstruir, desde un punto de vista lógico, enunciados de este tipo.

Una opción es interpretar el enunciado *default* "Las aves vuelan" como "Para todo x, *si* x es ave, entonces *normalmente* x vuela", es decir, empleando los cuantificadores del lenguaje de primer orden:

$\forall x$ (*si* ave(x), *entonces normalmente* vuela(x)).

Aquí, la expresión "normalmente" se refiere al condicional, que entonces se lo puede llamar un "condicional de normalidad" de la forma general

si p *entonces normalmente* q.

Este sería un tipo especial de condicional, es decir, un operador lógico que conecta enunciados, representable mediante el símbolo >, dando lugar a enunciados de la forma $p>q$. La idea es que si se determinaran las propiedades de este condicional especial se estaría analizando la naturaleza de la inferencia revocable.

Esta idea, de una manera ligeramente diferente, ya ha sido explotada (Delgrande, 1988). Su atractivo fundamental reside en que la revocabilidad recibe un tratamiento semántico análogo al de otros operadores que son extensiones de la lógica de primer orden (piénsese en la implicación estricta). Este tratamiento semántico se basa en la idea de mundo posible y en una función que selecciona de los mundos posibles aquellos que son “normales” (aquellos en los que se cumplen las condiciones de normalidad referidas por el enunciado *default*). Dicho informalmente, un condicional $p > q$ es verdadero a condición de que en los mundos normales en que p es verdadero, también q sea verdadero (Delgrande, 1988). Ciertamente, todos los principios del condicional revocable son válidos para el condicional clásico de la lógica de primer orden, representado por la flecha $\rightarrow$. Ahora bien, para el condicional > no vale la regla del *modus ponens*,

$(mp)\ A \rightarrow B, A \vdash B$

típica del condicional de la lógica de primer orden, pues, de otro modo, se estarían *deduciendo* enunciados que no pueden afirmarse de manera irrevocable, enunciados que pueden llegar a tener que ser descartados. Tómese nuevamente el ejemplo de Tweety:

- ave(Tweety) > vuela(Tweety)
- ave(Tweety)
  no puede deducirse
- vuela(Tweety).

Este enunciado sólo se podría inferir de una manera revocable.
Por lo tanto, el sistema lógico del condicional de normalidad > puede dar lugar a razonamientos que tengan como conclusión enunciados de la forma $p > q$, pero el consecuente de estos condicionales no podrá separarse de su respectivo antecedente. La revocabilidad quedará "encapsulada" en el condicional, sin dar lugar a inferencias revocables de enunciados. Dicho de otro modo, el sistema permite hacer inferencias (deductivas) *acerca* de *defaults*, pero no permite *usar* los *defaults* para hacer inferencias.[6]

Otra estrategia para analizar los enunciados *default* consiste en concentrarse no en el condicional, sino en la expresión "la mayoría de" y entenderlos como enunciados generales que contienen un tipo especial de cuantificador que representa la idea de "muchos", "la mayoría", etc. Esta interpretación se vincula con la idea de *razonamiento genérico*, es decir de razonamientos que se refieren a individuos genéricos o típicos, que representan la "normalidad", los casos típicos, de una clase de individuos (la típica ave, etc.). El enunciado *default* "Las aves vuelan" querrá decir "Toda ave típica vuela".

Sobre la base de la lógica clásica de primer orden, se puede introducir un nuevo símbolo lógico ∇ y construir un sistema $L\nabla$ para la lógica del razonamiento genérico (que es un caso de la llamada "lógica de ultrafiltros", véase Carnielli & Veloso, 1997). Algunas de las leyes para el cuantificador ∇ son:

---

[6] Observaciones adicionales pueden verse en Legris 1996.

(1) $\forall xA[x] \rightarrow \nabla xA[x]$;

(2) $\nabla xA[x] \rightarrow \exists xA[x]$;

(3) $\nabla xA[x]$ & $\nabla xB[x] \rightarrow \nabla x(A[x]\&B[x])$;

(4) $\neg\nabla xA[x] \rightarrow \nabla x\neg A[x]$.

(5) $\nabla xA[x] \vee \nabla x\neg A[x]$.

Pero nótese que no vale en general:

- $\nabla xA[x] \rightarrow A[a]$,

pues esto sólo valdrá, si a denota un elemento típico (es decir, si pertenece a un subconjunto del dominio de individuos, aquél cuyos elementos son "típicos" respecto de un predicado). Por ejemplo, no es aceptable la validez de la inferencia

- ∇x (ave(x) $\rightarrow$ vuela(x)) $\vdash$ ave(Tweety) $\rightarrow$ vuela(Tweety),

al no poder asegurarse que Tweety sea un objeto típico dentro de las aves.[7] Por lo demás, el sistema resultante es monótono. Por ejemplo, usando el axioma (2) y el *modus ponens* de la lógica clásica, se infiere del enunciado "∇x vuela(x)" el enunciado "$\exists x$ vuela(x)", y esta inferencia sigue siendo válida independientemente de cualquier enunciado que se añada a la premisa (incluso, por ejemplo, "¬vuela(Tweety)").

4. LA DEFINICIÓN DE OPERADORES LÓGICOS Y LA DEDUCCIÓN

En las dos perspectivas que se acaban de analizar, los sistemas resultantes son monótonos y los operadores introducidos para capturar la idea de revocabilidad pueden ser definidos de diferentes maneras, pero todas ellas responden a concepciones básicas de la lógica deductiva.

Los operadores lógicos, tanto los usuales de la lógica de primer orden (conjunción, negación, condicional, etc.) como los de las extensiones de la lógica de primer orden (lógicas modales, deónticas, temporales, etc.), aparecen en enunciados que se usan para hacer *afirmaciones* (categóricas). Por esta razón, se definen mediante las *condiciones suficientes* para afirmar enunciados que los incluyen. Estas condiciones pueden formularse como condiciones de verdad o afirmación. Se afirma un enunciado de la forma $A\&B$ si se afirman tanto A como B, y esta afirmación de $A\&B$ es *irrevocable*. Es decir, la afirmación de $A\&B$ tiene en la afirmación tanto de A como de B sus condiciones suficientes, independientemente de que estas condiciones se cumplan o no en un caso concreto: Si se dan las condiciones para afirmar A y las condiciones para afirmar B, entonces se dan las condiciones para afirmar $A\&B$.

Esto es equivalente con el hecho de que los operadores lógicos se definen en *contextos de deducción*, en los que valen las tres propiedades de reflexividad, transitividad y monotonía. Es decir, los operadores lógicos y la relación de deducción están esencialmente vinculados. Tal como se puede ver en cualquier ejemplo elemental, la determinación de la inferencia lógica depende de los operadores lógicos, que aparecen como *constantes* en la forma lógica. Dado un

---

[7] El sistema requiere alguna manera de representar individuos típicos.

razonamiento, su corrección deductiva dependerá de la corrección de su forma lógica, la que a su vez está determinada por la caracterización de los operadores lógicos que aparecen en él. La lógica no es otra cosa que la teoría de las constantes lógicas.

Ahora bien, el problema que subyace a los razonamientos revocables es que en ellos se afirman enunciados sobre la base de contar únicamente con *condiciones parciales* para hacer esa afirmación. Afirmar que Tweety vuela sobre la sola base de que la mayoría de las aves vuela y que Tweety es ave es afirmarlo sobre bases insuficientes (que *la mayoría* de las aves vuele no es *garantía* de que Tweety lo haga). Y puesto que toda deducción lógica válida resulta de principios que caracterizan los operadores lógicos, las condiciones parciales se retrotraerían a la definición de los operadores lógicos, lo que es inaceptable.

Tómese por caso afirmar una conjunción $A\&B$ sobre la base de una condición parcial, por ejemplo que se de únicamente A. En este caso, el completamiento de esas condiciones podría darse mediante $\neg B$ lo que conduciría a una inconsistencia, ya que de $A\&B$ puede inferirse B. Precisamente, dar las condiciones suficientes para afirmar un enunciado implica que ninguna circunstancia puede modificar la afirmación de ese enunciado.

Por lo demás, modificar alguna de estas condiciones conduciría a modificar la definición de los operadores lógicos: en un contexto no deductivo, los operadores lógicos usuales tendrían un significado diferente. Por ejemplo, si no se acepta la propiedad de monotonía no será válido el principio de la lógica clásica de primer orden

Si $\vdash B$,
entonces $\vdash A \to B$,

e igualmente no será válida la siguiente afirmación acerca de la relación de deducción
Si $A \vdash C$,
entonces $A\&B \vdash C$.[8]

Así pues, no parece recomendable tratar el problema de la revocabilidad al modo en que se han tratado las lógicas modales u otras lógicas extendidas, esto es, introduciendo nuevos operadores lógicos, sino que será mejor aceptar la lógica tal como está y modificar más bien la *aplicación* que se hace de la lógica en contextos más complejos. Es decir, la estrategia consistirá no en modificar la lógica, sino hacer modificaciones en la manera en que se da la información a partir de la cual se hacen las inferencias. Esto implica *agregar estructura* a la información.

5. LOS RAZONAMIENTOS REVOCABLES COMO RAZONAMIENTOS HIPOTÉTICOS

Tomando en cuenta los comentarios precedentes, una manera de tratar el problema del razonamiento revocable consiste en

[8] Un análisis más detallado de estos fenómenos pueden verse en Legris & Lerner 2000.

interpretar los enunciados *default* como "La mayoría de las aves vuelan" que aparecen en este tipo de razonamientos como *enunciados hipotéticos*, enunciados que se afirman de manera tentativa y cuya función es la de *justificar* parcialmente la afirmación de otros enunciados de un modo consistente con la información dada. En esta perspectiva, el carácter *default* de un enunciado no radica en algún operador lógico o en su forma, sino en una *cualificación* que recibe el enunciado respecto de una categorización del conocimiento. En efecto, el hecho de que un enunciado sea una hipótesis es un problema *pragmático* ligado al lugar que ocupa el enunciado en una estructura de conocimiento. El carácter de hipotético de un enunciado no es algo intrínseco al enunciado mismo, sino al hecho de cómo se lo usa en la obtención y el procesamiento de información.

Enunciados hipotéticos son esenciales para la formación de teorías estudiadas en la filosofía de la ciencia. En la formación de teorías, por ejemplo, se formulan hipótesis para *explicar* un fenómeno representado por uno o más enunciados y para *predecir* otros fenómenos semejantes. La formulación de hipótesis depende de ciertos requisitos pragmáticos y epistémicos, como, por ejemplo, su relevancia con el hecho a explicar, pero –sobre todo- las hipótesis se sostienen en la medida que no aparezcan nuevos datos que conduzcan a inconsistencias y lleven a descartarlas (este es el caso en que una hipótesis queda *refutada*). Procedimientos semejantes aparecen en la descripción de mecanismos refutatorios como parte de la "lógica de la investigación científica", desarrollados por Karl

Popper, o de los procedimientos abductivos basados principalmente en algunas ideas de Charles Peirce.

En el ejemplo de Tweety el enunciado “Las aves vuelan” funcionaría como una hipótesis que permite dar cuenta del fenómeno de que Tweety vuela, siempre y cuando no haya otros hechos que hagan inconsistente la justificación. La preservación de consistencia resulta una nota esencial de los razonamientos revocables que esta perspectiva desde el razonamiento hipotético destaca explícitamente. De acuerdo con esta idea, entonces, se tiene que en un razonamiento hipotético la conclusión debería verse como un enunciado a justificar o explicar sobre la base de ciertos datos afirmados como verdaderos y uno o más enunciados hipotéticos o *defaults*. A continuación se ofrece un primer esquema formal de esta reconstrucción de los razonamientos revocables (Poole, 1988).

Teniendo como base el lenguaje de primer orden (LPO), el marco formal tiene como base, en su caso más simple, una estructura $<H, F>$ que puede llamarse una *teoría default*, donde hay:

(i) un conjunto H de fórmulas abiertas de LPO que corresponden a los *defaults* o enunciados hipotéticos;

(ii) un conjunto consistente de fórmulas cerradas F de “hechos” o enunciados afirmados como verdaderos que representan datos dados que no son revisables.

A su vez, un *escenario* de <*H*, *F*> será una estructura <*D*, *F*>, donde *D* es un conjunto de instancias cerradas de elementos de *H*, tal que el *D* y *F* forman un conjunto consistente.

Sobre esta base un *razonamiento revocable* será una estructura que tendrá como conclusión una fórmula cerrada *A*, la cual se deducirá de *D* y *F*.

En otras palabras, dada una teoría *default* <*H*, *F*>, un razonamiento *default* con una fórmula cerrada *A* del LPO (por el momento atómico) como conclusión tiene como condición necesaria $D, F \vdash A$, siendo $\{D, F\}$ consistente y se dirá en ese caso que $H, F \mid\!\sim A$, donde $\mid\!\sim$ expresa una relación de inferencia revocable ("explicación" en Poole, 1988). Nótese la importancia esencial que tiene la consistencia para determinar que se da $\mid\!\sim$. Esta relación se caracteriza sobre la base de dos conceptos puramente *lógicos*: deducción y consistencia. No obstante, se presupone también la distinción *extralógica* entre hipótesis y hechos, distinción que lleva a cualificar fórmulas y a considerar un razonamiento revocable como una estructura *más compleja* que un razonamiento deductivo. El ejemplo de Tweety se reconstruiría como sigue:

$H = \{\text{ave}(x) \rightarrow \text{vuela}(x)\}$,

$F = \{\text{ave(Tweety)}\}$,

$D = \{\text{ave(Tweety)} \rightarrow \text{vuela(Tweety)}\}$.

$\{\text{ave(Tweety)} \rightarrow \text{vuela(Tweety)}, \text{ave(Tweety)}\}$ es consistente.

$\text{ave(Tweety)} \rightarrow \text{vuela(Tweety)}, \text{ave(Tweety)} \vdash \text{vuela (Tweety)}$.

Así, se sigue en esta primera aproximación que

$$\text{ave}(x) \rightarrow \text{vuela}(x),\ \text{ave(Tweety)}\ \mid\!\sim\ \text{vuela (Tweety)}.$$

Nótese que el razonamiento *default* contiene fórmulas *abiertas*; no puede considerarse un auténtico razonamiento de la lógica de primer orden (donde no se admiten fórmulas abiertas como premisas o conclusión). La relación de inferencia revocable $\mid\!\sim$ así caracterizada no cumple, como es de esperar, con la propiedad de monotonía. Claramente, si se agrega un enunciado que haga inconsistente al conjunto de premisas (que en el ejemplo anterior podría ser simplemente "¬vuela (Tweety)"), la relación de inferencia no se cumple.

6. HIPÓTESIS RIVALES Y EXTENSIONES

El agregado como una nueva hipótesis del enunciado "Los pingüinos no vuelan" y el hecho de que Tweety es un pingüino genera una situación diferente. Ahora se tiene:

H = {ave(x) → vuela(x), pingüino(x) → ¬vuela(x)},
F = {ave(Tweety), pingüino(Tweety)}.

Instanciando todas las hipótesis, resulta
D = {ave(Tweety) → vuela(Tweety), pingüino(Tweety) → ¬vuela(Tweety)},

que da lugar a un conjunto inconsistente de enunciados, de modo que no constituye un escenario y no puede decirse que "vuela(Tweety)" quede justificado por H y F. Los dos *defaults* aparecen como "hipótesis rivales" (tomando una terminología de la filosofía de la ciencia), es decir, hipótesis que no pueden adoptarse ambas de manera consistente, debiéndose rechazar una de ellas.

Ahora bien, si se adopta alternativamente uno u otro de los *defaults*, se pueden construir dos escenarios diferentes consistentes

D_1 = {ave(Tweety) $\rightarrow$ vuela(Tweety)}

D_2 = {pingüino(Tweety) $\rightarrow$ $\neg$vuela(Tweety)}

respecto de los cuales quedan justificados respectivamente los enunciados "Tweety vuela" y "Tweety no vuela", si bien no puede afirmarse que ambos enunciados queden justificados *en sentido estricto* por el conjunto de hipótesis. Se trata en este caso más bien de un apoyo o justificación *más débil* de cada uno de estos enunciados, apoyo del tipo "existe alguna *situación* en la que puede inferirse que Tweety vuela", por ejemplo. Es decir, en este caso se dan dos *situaciones* diferentes y además mutuamente excluyentes que ofrecen respaldo respectivamente a enunciados contradictorios entre sí.

En otras palabras, uno puede elegir dos conceptos diferentes de inferencia revocable, uno más fuerte y otro más débil:

(a) $H, F \mid\sim A$, si para *todo* escenario $\langle D, F\rangle$ de $\langle H, F\rangle$, $D, F \vdash A$.

(b) $H, F \mid\sim A$, si para *algún* escenario $\langle D, F\rangle$ de $\langle H, F\rangle$, $D, F \vdash A$.

7. PREFERENCIA DE HIPÓTESIS

Surge aquí, entonces, la pregunta de si alguna de las dos situaciones es preferible o, dicho de otro modo, si tiene un apoyo epistémico mayor que la otra, es decir, la pregunta de cuál de las dos es más aceptable epistémicamente. Esto conduce a un punto de vista diferente, en el cual se exige un *orden de preferencia epistémica* entre las diferentes justificaciones. Este orden estará determinado, en definitiva, por el orden de preferencia que se les asigne a conjuntos de hipótesis, cuyos miembros tengan un mismo nivel de preferencia.

(Siguiendo a Brewka, 1991), un procedimiento podría basarse en la idea de determinar una relación de orden >, que debería ser *prima facie* un orden estricto y total entre *conjuntos* de hipótesis (los elementos de ese conjunto tendrían el mismo grado de confiabilidad epistémica).[9] Esto iría junto con el abandono de la diferenciación entre hipótesis y "hechos", considerando a *toda* afirmación como hipotética. En efecto, lo que distingue epistémicamente a los *defaults* de los hechos es su mayor "arraigo" epistémico. Algunas de ellas, no obstante tendrán un valor epistémico mayor que otras (serán epistémicamente más confiable que otras, o estarán más justificadas que otras) y, por lo tanto, serán preferidas. Frente a una inconsistencia, se elige el conjunto de hipótesis epistémicamente más confiable. Así, una teoría *default* T será ahora una secuencia de hipótesis H_1, ..., H_n ordenada por la relación >.

[9] El signo ">" tiene, por cierto, un significado diferente que en la sección 3, donde se lo usaba para representar condicionales revocables. La diferencia de contexto hace que no sea necesario usar signos diferentes.

Volviendo al ejemplo precedente, supóngase que la hipótesis "pingüino(x) $\rightarrow$ ¬vuela(x)" es más confiable epistémicamente que la hipótesis "ave(x) $\rightarrow$ vuela(x)", es decir, se tiene

{pingüino(x) $\rightarrow$ ¬vuela(x)} > {ave(x) $\rightarrow$ vuela(x)},

entonces el enunciado "Tweety no vuela" estaría *más justificado* que el enunciado "Tweety vuela" y sería el enunciado a inferir revocablemente. Por lo tanto, pueden sostenerse como premisas ambas hipótesis, ya que tienen un lugar diferente en el orden de preferencia epistémica.

Esta perspectiva ofrece una solución al problema de la monotonía: a las hipótesis que constituyen las premisas de un razonamiento se les puede agregar cualquier conjunto de hipótesis a condición de que sea menos preferible. Así, no se altera la inferencia del enunciado. Todo esto puede formularse con más precisión por medio de la siguiente definición de la relación de inferencia revocable $\mid\!\sim$.

Sea T un conjunto de conjuntos de hipótesis, entonces $T \mid\!\sim A$
si y sólo si
existe un subconjunto S de T, tal que
(i) no $S \vdash \perp$,
(ii) $S \vdash A$, y
(iii) para todo H_i ε $T - S$, si S, $H_i \vdash \perp$, entonces para todo H_j ε S, $H_j > Hi$
(el signo $\perp$ representa un enunciado contradictorio cualquiera).

Para la relación así definida vale una versión limitada de monotonía que puede denominarse *monotonía preferencial*, caracterizada del siguiente modo

(MP) Si $T \mid\sim A$,

entonces $T,\ H \mid\sim A$, a condición de que para todo $H_i \in T$, $H_i > H$.

Para resumir, dos cosas pueden observarse claramente aquí. En primer lugar, en estos razonamientos la *conservación de consistencia* es un objetivo fundamental. En segundo lugar, este caso muestra que, en general, los razonamientos no pueden caracterizarse exclusivamente en términos de una relación entre enunciados, sino que es necesario considerar *estructuras* de enunciados.

Cabe mencionar, finalmente, un problema que presentan tanto este como los demás formalismos que se han propuesto para reconstruir el concepto de inferencia revocable. Estos sistemas no son tratables computacionalmente. De hecho, no son ni siquiera semidecidibles, es decir, no existe un algoritmo para inferir revocablemente un enunciado a partir de una estructura como la recién presentada. Esto se debe a que se basan en el concepto de consistencia, que, en general, no es semidecidible. Pese a este resultado, negativo, es posible construir procedimientos de demostración para casos particulares, o se pueden introducir ciertas limitaciones en la capacidad expresiva del lenguaje, de modo de hacer al sistema computacionalmente tratable.[10]

REFERENCIAS

[1] Brewka, Gerhard. 1991. *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge, et al., Cambridge University Press.

[2] Carnielli, Walter A. & Paulo A. S. Veloso. 1997. "Ultrafilter Logic and Generic Reasoning". En *Computational Logic and Proof Theory*, comp. por G. Gottlob, A. Leitsch y D. Mundici. Berlín, Springer, pp. 34-53.

[3] Delgrande, J.P. 1988. "An Approach to Default Reasoning Based on a First Order Conditional Logic: Revised Report". En *Artificial Intelligence* 36, pp. 63-90.

[4] Legris, Javier. 1996. "Diferentes criterios para analizar la inferencia revocable: observaciones sobre el punto de vista lógico". En *Jornadas de Epistemología de las Ciencias Económicas 1995*, comp. por Pablo S. García, Gustavo Marqués y Eduardo R. Scarano, Buenos Aires, Secretaria de Investigación y Doctorado-Facultad de Ciencias Económicas, pp. 129-135.

[5] Legris, Javier & Lerner, Silvia. 2000. "Observaciones sobre la formalización de la inferencia revocable". En *Actas de las V Jornadas de Epistemología de las Ciencias Económicas 1999*, comp. por Javier Legris. Buenos Aires, Economizarte-Imprenta de la Facultad de Ciencias Económicas, pp. 184-189.

---

10 Sugerencias al respecto pueden encontrarse en Lombardi 1999.

[6] Lombardi, Carlos. 1999. "Ideas acerca de una teoría de la demostración para una reconstrucción de razonamientos revocables". En *Epistemología e Historia de la Ciencia* 5, pp. 242-249.

[7] Poole, David. 1988. "A Logical Framework for Default Reasoning". *Artificial Intelligence* 36, pp. 27-47.

[8] Poole, David, Alan Mackworth & Randy Goebel. 1998. *Computational Intelligence. A Logical Approach*. Nueva York-Oxford, Oxford University Press.

[9] Reiter, Raymond. 1987. "Nonmonotonic Reasoning". *Ann. Rev. Comput. Sci.* 2, pp. 147-186.

[10] Simon, Herbert. 1981. *The Sciences of The Artificial*. 2da. ed., Cambridge (Mass.)-London, The MIT Press.